

EP99/3031



EPO - DG 1

20 07. 1999



## Bescheinigung

REC'D 27 JUL 1999	
WIPO	PCT

Die VOLKSWAGEN Aktiengesellschaft in Wolfsburg/Deutschland hat eine Patentanmeldung unter der Bezeichnung

"Verfahren und Einrichtung zum Betrieb von sprachunterstützten Systemen in Kraftfahrzeugen"

am 18. Juni 1998 beim Deutschen Patent- und Markenamt eingereicht und erklärt, daß sie dafür die Innere Priorität der Anmeldung in der Bundesrepublik Deutschland vom 6. Mai 1998, Aktenzeichen 198 20 000.5, in Anspruch nimmt.

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

Die Anmeldung hat im Deutschen Patent- und Markenamt vorläufig das Symbol G 10 L 15/20 der Internationalen Patentklassifikation erhalten.

München, den 30. Juni 1999

**Deutsches Patent- und Markenamt**

**Der Präsident**

Im Auftrag

Heiß

Aktenzeichen: 198 27 134.4

**VOLKSWAGEN**

K 6970/1770-hk-ch



## **Verfahren und Einrichtung zum Betrieb von sprachunterstützten Systemen in Kraftfahrzeugen**

Die Erfindung betrifft ein Verfahren sowie eine Einrichtung zum Betrieb von sprachunterstützten Systemen, wie Kommunikations- und/oder Sprech- /Gegensprech-einrichtungen in Kraftfahrzeugen, bei welchem bzw. bei welcher über eine Mehrfachmikrofonanordnung Sprachsignale aufgenommen und an mindestens einen Lautsprecher weitergegeben werden, gemäß Oberbegriff der Patentansprüche 1 und 7.

Verfahren dieser Art werden in Kraftfahrzeugen zum einen zum sprachunterstützten Gegensprechbetrieb eingesetzt, oder aber auch zur Unterstützung von spracheingabegesteuerten elektronischen oder elektrischen Baugruppen. Die grundsätzliche Problematik hierbei ist, daß im Kraftfahrzeug je nach Betriebszustand eine entsprechende Geräuschkulisse vorhanden ist. Diese überdeckt die Sprachbefehle. Sprech- und Gegensprechanlagen in Kraftfahrzeugen sind überwiegend bei großen Fahrzeugen, Minibussen und dergleichen vorteilhaft. Sie können jedoch auch bei normalen Personenkraftwagen eingesetzt werden. Bei der Verwendung von sprachgesteuerten Eingabeeinheiten für elektrische Komponenten im Fahrzeug ist die Unterdrückung der Geräuschkulisse bzw. das Herausfiltern des Sprachbefehles noch von besonderer Bedeutung.

So ist aus der EP 0078014 B 1 eine Spracherkennungseinrichtung für ein Kraftfahrzeug bekannt, bei welchem in das Verstärkersystem der Spracherkennungseinrichtung über Sensoren gemeldet bzw. eingespeist wird, ob der Motor in Betrieb ist und/oder sich das Fahrzeug bewegt. Danach richtet sich sodann eine Pegalbeeinflussung mit der versucht wird, den Sprachbefehl aus der Geräuschkulisse herauszufiltern.

Aus der DE 3742929 C 1 ist eine Anordnung mit zwei Mikrofonen bekannt, wobei eines der Mikrofone am Mund der Bedienperson angeordnet ist und ein anderes in der Nähe, jedoch

zur Aufnahme des Körperschalles. Beide Mikrofonsignale werden getriggert derart, daß Körperschall vom Gesamtschall subtrahierbar ist.

Aus der DE 19705471 A 1 ist bekannt, eine Spracherkennung mit Hilfe einer Transversalfilterung zu unterstützen. Hierbei wird eine Frequenzanalyse vorgenommen, die jedoch lediglich zu dem Zweck der Sprachbefehlserkennung dient. Es findet hierbei keine Nebengeräuschkompensation statt.

Aus der WO 97/34290 ist eine Filterung bekannt, bei der periodische Störsignale ausgefiltert werden, in dem deren Periode ermittelt und mittels Generator herausinterferiert wird, so daß das Sprachsignal übrig bleibt.

Aus der DE 4106405C 2 ist ein Verfahren bekannt, bei dem eine Geräuschsubtraktion vom Sprachsignal erfolgt, wobei eine Mehrzahl von Mikrofonen verwendet wird.

Aus der DE 3925589 A 1 ist die Verwendung einer Mehrfachmikrofonanordnung bekannt, wobei bei Anwendung im Kraftfahrzeug eines der Mikrofone im Motorraum und ein weiteres im Fahrgastraum angeordnet ist. Sodann erfolgt eine Subtraktion beider Signale. Nachteilig ist hierbei, daß lediglich das Motorgeräusch bzw. das eigentliche Betriebsgeräusch des Fahrzeuges selbst vom Gesamtsignal im Fahrgastraum abgezogen wird. Spezifische Nebengeräusche werden hierbei unberücksichtigt gelassen. Ebenso fehlt eine Rückkopplungsunterdrückung, die eine besondere Problematik darstellt. Überall dort, wo Mikrofone und Lautsprecher in akustisch ankoppelbarer Nähe angeordnet sind, kommt es vor, daß das am Lautsprecher ausgekoppelte akustische Signal wiederum in das Mikrofon rückeinspeist. Es kommt zu einer sogenannten Rückkopplung und einer darauf folgenden Übersteuerung.

Der Erfindung liegt somit die Aufgabe zugrunde, ein Verfahren sowie eine Einrichtung der gattungsgemäßen Art dahingehend weiterzubilden, daß Rückkopplungen und Instabilitäten die bei Anordnung mehrerer Mikrofone und Lautsprecher auftreten zu unterdrücken.

Die gestellte Aufgabe ist bei einem Verfahren der gattungsgemäßen Art erfindungsgemäß durch die kennzeichnenden Merkmale des Patentanspruches 1 gelöst.

Weitere vorteilhafte Ausgestaltungen des Verfahrens sind in den Ansprüchen 2 bis 5 angegeben.

Hinsichtlich einer Einrichtung der gattungsgemäßen Art ist die gestellte Aufgabe erfindungsgemäß durch die kennzeichnenden Merkmale des Patentanspruches 6 gelöst. Weitere vorteilhafte Ausgestaltungen der einrichtungsgemäßen Erfindung sind in den übrigen Ansprüchen angegeben.

Die Erfindung geht sowohl hinsichtlich des Verfahrens als auch der Einrichtung von einer Kommunikations- und/oder Sprech- /Gegensprecheinrichtung in Kraftfahrzeugen aus. Es ist auch bekannt, hierbei eine Mehrfachmikrofonanordnung anzuordnen, darüber hinaus Sprach- als auch Geräuschsignale aufzunehmen und vom Gesamtsignal die Geräuschsignale wiederum zu subtrahieren, so daß das Sprachsignal gefiltert übrig bleibt.

Gemäß der gestellten Aufgabe besteht der Kern der Erfindung darin, daß das jeweilige Mikrofonsignal zunächst um einen kleinen Betrag  $\Delta F$  frequenzverschoben wird, und nachfolgend erst auf den oder die Lautsprecher oder auf die Eingabe einer sprachgesteuerten Einrichtung gegeben wird. Die erfindungsgemäße Frequenzverschiebung, die hierbei an definierter Stelle vorgenommen wird und nicht willkürlich ist, unterstützt zum einen die Filterung und zum anderen werden Rückkopplungen, also auch das Echosignal ausgekoppelt.

Da Rückkopplungen ohne die besagte erfindungsgemäße Frequenzverschiebung nichts weiter sind als das rückgekoppelt verstärkte Sprachsignal, können mit Mitteln und Vorgehensweisen aus dem zitierten Stand der Technik solche Rückkopplungen nicht eliminiert werden. Dies aus dem besagten Grund, weil Einrichtungen der bekannten Art lediglich das Sprachsignal vom Geräuschsignal separieren und das rückgekoppelte Signal als Sprachsignal und nicht als Geräuschsignal identifizieren. Dadurch sind die besagten Rückkopplungen mit Hilfe der im Stand der Technik bekannten Mittel nicht, oder nicht gleichzeitig beherrschbar.

Demgegenüber wird jedoch durch das erfindungsgemäße Verfahren sowie durch erfindungsgemäße Einrichtung, die sich auf die Verschaltung der einzelnen Elemente miteinander bezieht werden auf elegante Weise Rückkopplungseffekte eliminiert.

Da die Rückkopplung als solches ursächlich immer dann auftritt, wenn Mikrofonort und Lautsprecherort dicht beieinander liegen, was in Kraftfahrzeugen zwanghaft der Fall ist, kommt der Eliminierung dieser Rückkopplung im genannten Anwendungsfall ganz

erhebliche Bedeutung zu. Dies gilt nicht nur im Falle des Gegensprechbetriebes, bei dem elektroakustische Rückkopplungen für die Insassen unangenehm sind, sondern besondere Bedeutung kommt auch beim Einsatz sprachgesteuerter Eingabeschnittstellen von elektrischen bzw. elektronischen Bauteilen am Kraftfahrzeug zu. Dies gilt nur dann, wenn die gesamte Anordnung im Fahrzeug sowohl Mikrofone als auch Lautsprecher umfaßt, und hierüber auch die sprachgesteuerte Eingabe an elektrische Geräte erfolgt. Rückkopplungen und daraus resultierende Übersteuerungen können selbst bei intelligenten Eingabeschnittstellen erhebliche Fehlfunktionen und Mißdeutungen des Sprachbefehles hervorrufen. Je nach Einsatzfall stellt dies auch ein Sicherheitsrisiko dar. Wahlweise kann die gleichzeitige Geräuscheminierung zusätzlich, d.h. gleichzeitig vorgenommen werden.

Die Erfindung ist in der Zeichnung dargestellt und nachfolgend näher beschrieben.

Die Abbildung zeigt den prinzipiellen Aufbau sowie auch die Funktionsweise, so daß aus der Abbildung selbst sowohl die Verfahrensmaßnahmen, als auch die Verschaltung der einzelnen einrichtungsgemäßen Elemente zueinander in ihrer logischen Gesamtheit erkennbar ist, bzw. sind.

In diesem dargestellten Ausführungsbeispiel der Erfindung ist der Fahrzeuginnenraum in zwei Teilräume unterteilt, nämlich vorne und hinten.

Im vorderen Teil befindet sich ein Mikrofon M 1 und ein Lautsprecher L 2 .

Das Mikrofon M 1 nimmt das dortige Sprachsignal und ggf. Geräuschsignale auf. Das Geräuschsignal besteht dabei aus der sich im Betrieb des Fahrzeuges ergebenden Geräuschkulisse im Fahrgastraum. Dies können Motorengeräusche, Windgeräusche sowie Abrollgeräusche aber auch akustische Echosignale aus dem anderen Teilraum und dergleichen mehr sein. Das an M 1 enthaltene Summensignal aus Sprache- und Geräuschkulisse wird einem ersten Summationspunkt S 1 zugeführt. Diesem Summationspunkt wird dann ebenfalls ein entsprechend aufbereitetes Signal aus einem akustischen Modell AM 1 vorne, zugeführt. Das im akustischen Modell AM 1 generierte Subtraktionssignal entstammt in diesem Ausführungsbeispiel aus dem im hinteren Teil des Fahrzeuges erhaltenen, und bereits frequenzverschobenen Signal. Dadurch, daß dieses von M 2 kommende und in F 2 frequenzverschobene Signal, welches dem hinteren Teilraum der Fahrgastzelle entstammt, über AM 1 signaltechnisch auch vorne berücksichtigt wird, wird die im hinteren Teil des Fahrzeuges generierte und nach vorne, in den vorderen Teil der Fahrgastzelle akustisch transportierte Teil, welcher auch von M 1 registriert wird,

am Summationspunkt S 1 wiederum subtrahiert. D.h., durch die Einrichtung AM 1 wird der hintere Teilraum der Fahrgastzelle vom vorderen Teilraum der Fahrgastzelle akustisch getrennt. D.h., zunächst wird in M 1 das gesamte wahrnehmbare akustische Signal eingespeist, und am Summationspunkt S 1 zunächst das Echo vom hinteren Teilraum der Fahrgastzelle subtrahiert. Das so erhaltene originäre Signal von M 1 aus dem vorderen Teilraum der Fahrgastzelle wird sodann einer Frequenzverschiebeeinrichtung F 1 zugeführt und um einen Betrag  $\Delta F$ , beispielsweise 5 Hz, verschoben. Das so erhaltene Ausgangssignal von M 1 wird sodann dem Lautsprecher L1 des hinteren Teilraumes der Fahrgastzelle zugeführt und zum anderen gleichzeitig auch wiederum auf die gleiche Weise in die Einrichtung AM 2 eingespeist. AM 2 repräsentiert dabei wieder das akustische Modell für den hinteren Teilraum der Fahrgastzelle. Die Übermittlung einer Sprachnachricht vom hinteren Teilraum der Fahrgastzelle über M 2 zum vorderen Teilraum der Fahrgastzelle über L 2 erfolgt in analoger Weise. D.h., das Mikrofon M 2 registriert die Sprachnachricht samt Geräuschkulisse im hinteren Teilraum der Fahrgastzelle und übermittelt sie an den Summationspunkt S 2, an welchem das über M 1 aufgenommene akustische Gesamtsignal, d.h., das Echo sowie Nebengeräusche, subtrahiert wird. Das so wiederum erstellte echofreie Signal von dem Mikrofon M 2 wird sodann ebenfalls einer Frequenzverschiebeeinrichtung F 2 zugeführt, die wiederum eine Frequenzverschiebung um einen Betrag  $\Delta F$  vornimmt. Am Ausgang dieser Frequenzverschiebungseinrichtung F 2 wird das Ergebnis bzw. das so aufbereitete Signal wiederum dem vorderen Teilraum der Fahrgastzelle, nämlich dem dort positionierten Lautsprecher L 2 zugeführt. Die Frequenzverschiebung für die Übermittlung von vorne nach hinten kann auch von derjenigen Frequenzverschiebung von hinten nach vorne unterschieden sein.

Insgesamt ergibt sich ein geschlossenes rückkopplungsfreies System. Dabei ist die Verschiebung der Frequenz ein wesentliches Merkmal, und durch das Zusammenwirken mit der Verschaltung über die akustischen Modelle AM 1 und AM 2 ist eine Echoeliminierung vom vorderen zum hinteren Teilraum und umgekehrt gegeben.

Es ist jedoch auch möglich, daß zusätzlich zur Echounterdrückung und Rückkopplungseliminierung auch eine Geräuschsignalsubtraktion hinzukommt. Dies kann in geeigneter Weise im jeweiligen akustischen Modell AM1 und AM 2 mitberücksichtigt werden. Die weitergehenden, hierzu notwendigen Komponenten, wie Geräuschsignalmikrofone, sind dabei nicht weiter dargestellt.

Somit läßt sich also sagen, daß jedes akustische Eingangssignal von M 1 sowie auch von M 2, bevor es weitergehend aufbereitet und den Lautsprechern L 2 bzw. L 1 zugeführt wird, das aus Echo und übrigen Geräuschen zusammengesetzte Gesamtgeräuschkulissensignal subtrahiert wird. Es findet somit nicht nur eine akustische Abkopplung zwischen vorderem und hinterem Teilraum der Fahrgastzelle statt, sondern auch die übrigen Geräuschsignale werden quasi in ein und demselben Aktionsschritt mitkompensiert bzw. subtrahiert.

K 6970/1770-hk-ch

## PATENTANSPRÜCHE

1. Verfahren zum Betrieb von sprachunterstützten Systemen, wie Kommunikations- und/oder Gegensprecheinrichtungen in Kraftfahrzeugen, bei welchem über eine Mehrfachmikrofonanordnung Sprachsignale aufgenommen und an mindestens einen Lautsprecher weitergegeben werden, dadurch gekennzeichnet, daß das Sprachsignal bzw. das Sprachsignalspektrum zunächst um einen Betrag  $\Delta F$  frequenzverschoben wird und nachfolgend erst auf den oder die Lautsprecher oder auf die Eingabe einer sprachgesteuerten Einrichtung gegeben wird.
2. Verfahren zum Betrieb von sprachunterstützten Systemen, nach Anspruch 1, dadurch gekennzeichnet, daß jeweils vor Übermittlung des Signales an die Lautsprecher das Echo aus dem lautsprecherbezogenen Umfeld vom Signal des hierfür betätigten mikrofonbezogenen Umfeldes subtrahiert wird.
3. Verfahren zum Betrieb von sprachunterstützten Systemen, nach Anspruch 1 oder 2, dadurch gekennzeichnet, daß bei Verwendung einer Mehrzahl von Mikrofonen jedes akustische Aufnahmesignal eines jeden Mikrophones nach Subtraktion des jeweils lautsprecherbezogenen Umfeldes bzw. des dort generierten Geräuschsignales, um  $\Delta F$  frequenzverschoben wird.
4. Verfahren zum Betrieb von sprachunterstützten Systemen, nach einem oder mehreren der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß zur akustischen Ankopplung bzw. Subtraktion der Gesamtgeräuschkulissensignale ein beliebiges akustisches Modell aus den aufgenommenen Gesamtsignalen gebildet wird und signaltechnisch zwischen jeweiligem Mikrofon und jeweiliger

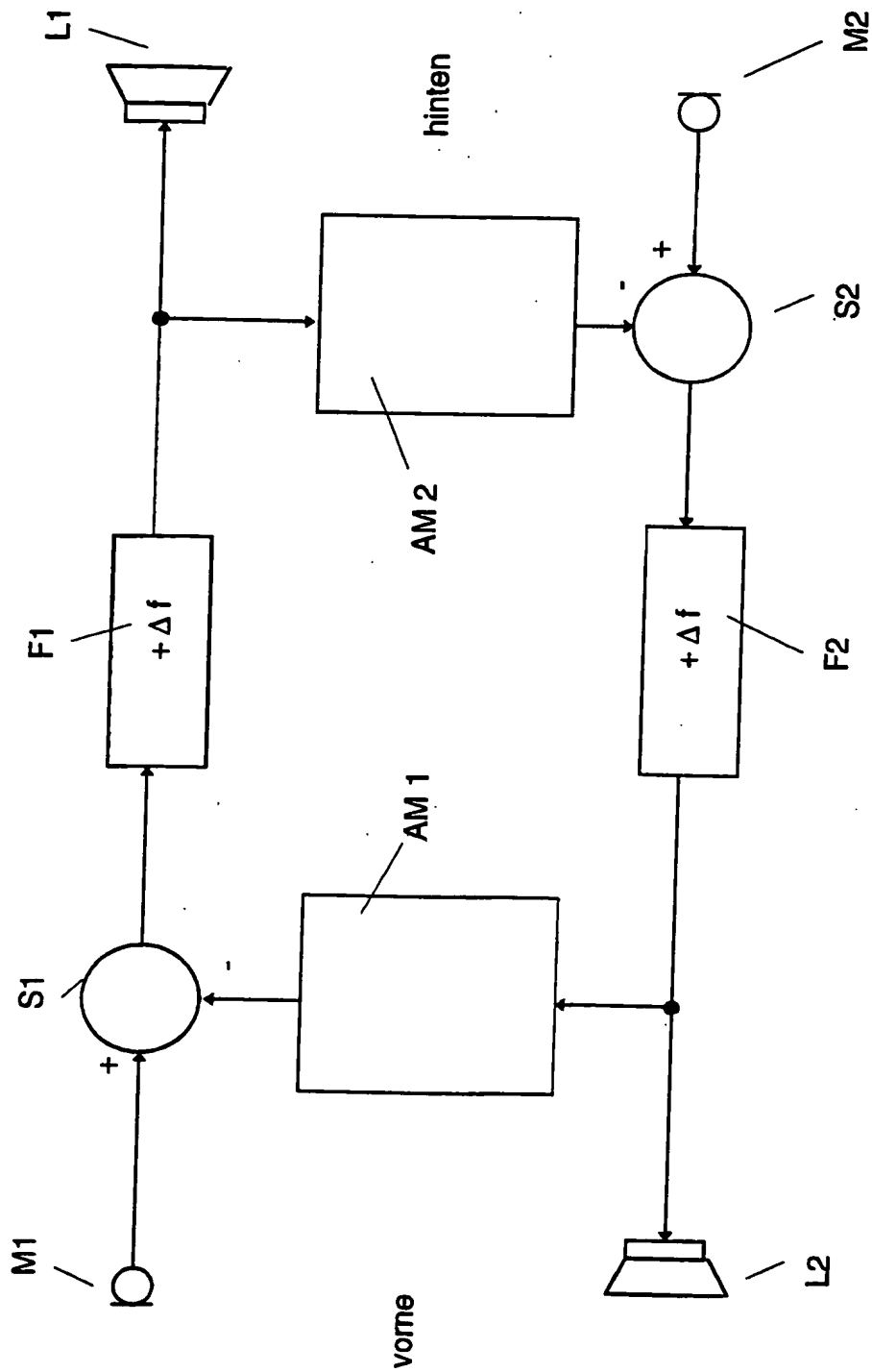


Frequenzverschiebung auf einen jeweiligen Summationspunkt zur Subtraktion aufgegeben wird.

5. Verfahren zum Betrieb von sprachunterstützten Systemen nach Anspruch 4, dadurch gekennzeichnet, daß der Fahrgastraum des Fahrzeuges in mindestens zwei akustisch Teilräume aufgeteilt ist, derart, daß in jedem Teilraum zumindestens ein Mikrofonort als auch zumindestens ein Lautsprecherort vorhanden ist, daß zwischen Mikrofonort des einen Teilraumes und dem Lautsprecherort des anderen Teilraumes die besagte Frequenzverschiebung  $\Delta F$  erfolgt und zwischen den Lautsprecherorten und Mikrofonorten des einen Teilraumes und zwischen den Lautsprecherorten und Mikrofonorten des anderen Teilraumes die besagten akustischen Modelle angewendet werden, so daß signaltechnisch ein geschlossener elektroakustischer Regelkreis entsteht.
6. Verfahren zum Betrieb von sprachunterstützten Systemen nach einem oder mehreren der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß über die besagten akustischen Modelle nicht nur die Sprach- und/oder Geräuschsignale der unterschiedlichen Teilräume in der Fahrgastzelle berücksichtigt werden, sondern zusätzlich ermittelte im gesamten Umfeld bestehende Geräusche mitberücksichtigt und vom Gesamtschallsignal subtrahiert werden, so daß im wesentlichen das Sprachsignal übrig bleibt.
7. Einrichtung zum Betrieb von sprachunterstützten Systemen, wie Kommunikations- und/oder Gegensprecheinrichtungen in Kraftfahrzeugen, mit einer Mehrzahl von Mikrofonen und Lautsprechern, sowie mit Mitteln zur Übertragung von Sprachnachrichten oder Sprachbefehlen dadurch gekennzeichnet, daß der Fahrgastraum im Kraftfahrzeug in mindestens zwei ggf. offene Teilbereiche (vorne, hinten) mit jeweils mindestens einem Mikrofon (M1, M2) und mindestens einem Lautsprecher (L1, L2) unterteilt ist, daß die besagten Mittel auch Frequenzverschiebeeinrichtungen (F1, F2) umfassen, welche zwischen jeweils einem der Mikrofone (M1, M2) und dem im jeweils anderen Teilbereich (vorne, hinten) befindlichen Lautsprecher geschaltet ist, und daß das jeweils resultierende Lautsprechersignal parallel abgreifbar und dem jeweils im selben Teilbereich

angeordneten Mikrofonsignal über Summationspunkte (S1, S2) subtraktiv überlagerbar ist.

8. Einrichtung zum Betrieb von sprachunterstützten Systemen nach Anspruch 7, dadurch gekennzeichnet, daß zwischen Parallelabgriff des jeweiligen Lautsprechersignales und dem jeweiligen Summationspunkt (S1, S2) Mittel (AM1, AM2) vorgesehen sind, über welche sogenannte akustische Modelle generierbar sind, die das jeweilige Lautsprechersignal beeinflussen/nachbearbeiten und das Ergebnissignal aus (AM1) und (AM2) auf den jeweiligen Summationspunkt schaltbar ist.
9. Einrichtung zum Betrieb von sprachunterstützten Systemen nach Anspruch 8, dadurch gekennzeichnet, daß die akustischen Modelle (AM1, AM2) Mittel zur Geräuschmustererkennung enthalten, die zur Separierung von Motor-/Fahrgeräuschen von sprachgenerierten akustischen Signalen, sowie zur Separierung primär sprachgenerierter Signale von rückgekoppelten Echosignalen dienen.



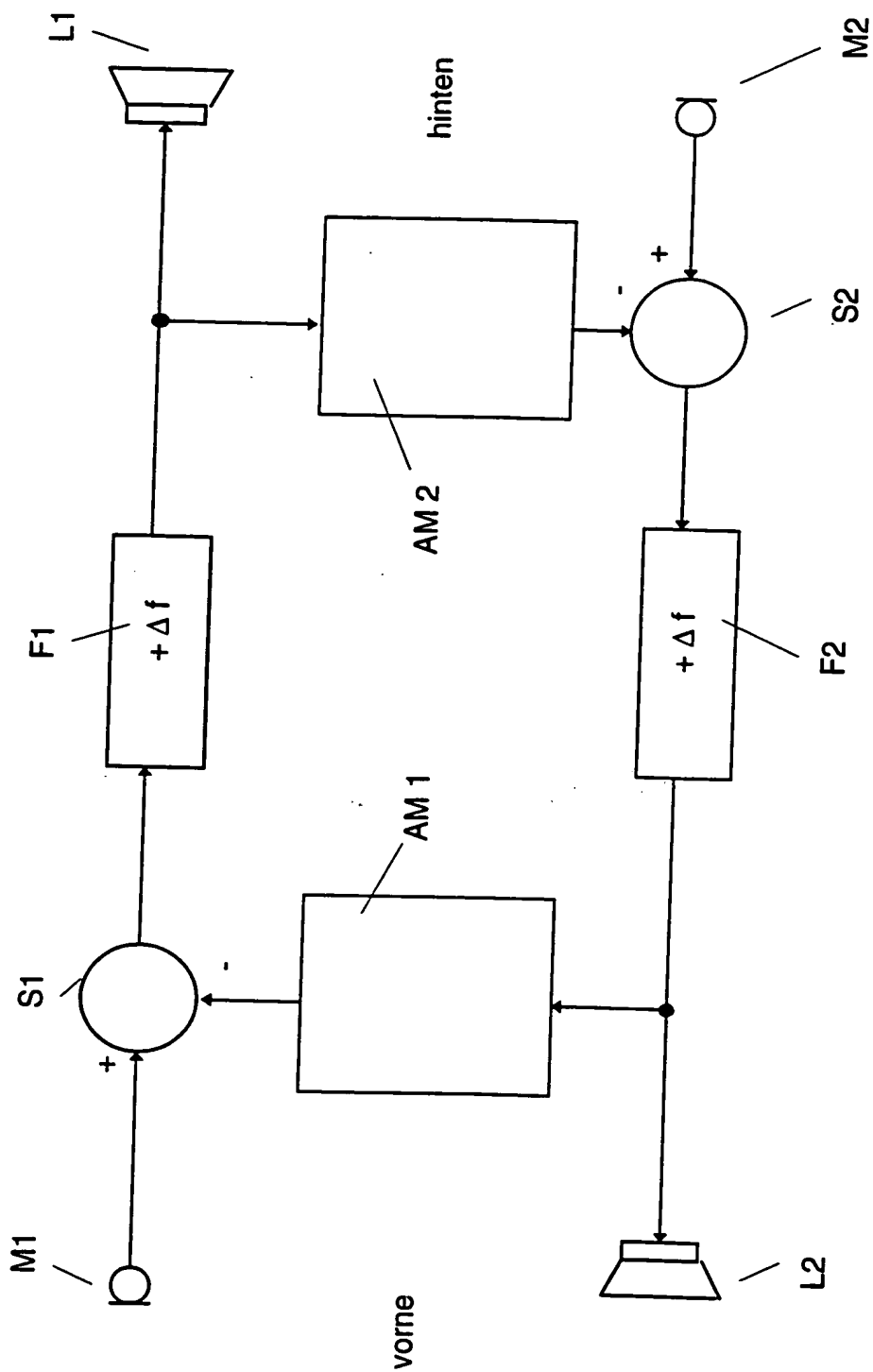
K 6970/1770-hk-ch

## ZUSAMMENFASSUNG

### Verfahren und Einrichtung zum Betrieb von sprachunterstützten Systemen in Kraftfahrzeugen

Die Erfindung betrifft ein Verfahren sowie eine Einrichtung zum Betrieb von sprachunterstützten Systemen, wie Kommunikations- und/oder Sprech/Gegensprecheinrichtungen in Kraftfahrzeugen, bei welchem bei einer Mehrfachmikrofonanordnung Sprachsignale aufgenommen und an mindestens einen Lautsprecher weitergegeben wird. Um bei einem Verfahren sowie einer Einrichtung dieser Art zu gewährleisten, daß Rückkopplungen eliminiert werden, ist erfindungsgemäß vorgeschlagen, daß das Sprachsignal bzw. das Sprachsignalspektrum zunächst um einen kleinen Betrag  $\Delta F$  frequenzverschoben wird und nachfolgend erst auf den oder die Lautsprecher oder auf die Eingabe einer sprachgesteuerten Einrichtung gegeben wird.

(Fig.)



**THIS PAGE BLANK (USPTO)**